Applied Mathematics and Computation 220 (2013) 239–245

Contents lists available at SciVerse ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

New step lengths in projection method for variational inequality problems

Yunda Dong *, Xue Zhang

School of Mathematics and Statistics, Zhengzhou University, 450001 Zhengzhou, PR China

ARTICLE INFO

Keywords:
Variational inequality
Projection method
Strong monotonicity
Lipschitz continuity
Self-adaptive step length

ABSTRACT

In this paper, we consider projection method for variational inequality problems. First, we give a new modification of recently proposed self-adaptive step length rules, with possibly minimal parameters. Then, the resulting self-adaptive projection method is proven to converge globally at a R-linear rate provided that the underlying mapping is strongly monotone and Lipschitz continuous. Preliminary numerical results confirm its flexibility and effectiveness.

1. Introduction

Let $F : R^n \to R^n$ be a continuous mapping and C be a nonempty closed convex subset of R^n. A classical variational inequality problem is to find an x^* in C such that

$$\langle x - x^*, \boldsymbol{F}(x^*)\rangle \geqslant 0, \quad \forall x \in C, \tag{1}$$

where $\langle \cdot, \cdot \rangle$ stands for the usual inner product, and the induced norm is defined by $\|x\| = \sqrt{\langle x, x\rangle}$ for any given x in R^n.

In this paper, we study the projection method [1] for solving the variational inequality problem above. For any given starting point $x^0 \in R^n$, it generates a sequence of the iterates $\{x^k\}$ by implementing the following recursion:

$$x^{k+1} = P_C[x^k - \beta_k F(x^k)], \tag{2}$$

where $\beta_k > 0$ is a step length, and P_C stands for the usual projection onto the set C. When F is the gradient of some real differentiable function f in R^n, it reduces to a gradient projection method of Goldstein [2] and of Levitin and Polyak [3]. See [4–8] for related discussions.

Such a projection method is known to converge globally provided that F is L-Lipschitz continuous and α-strongly monotone (see the definitions below), and the involved step length satisfies

$$0 < \beta_L \leqslant \beta_k \leqslant \beta_U < 2\alpha/L^2.$$

However, such step length is not implementable because both Lipschitz constant and strong monotonicity modulus are rather difficult to evaluate in practice.

To resolve such a difficulty, He et al. [9] proposed a practical step length rule of the projection method. Specifically speaking, given a summable sequence $\{\tau_k\}$ of positive numbers. Choose the parameters $\delta \in (0,1),\ \mu \in (0.5,1)$ and $\beta_0 > 0$. Choose $x^0 \in C$. Set $\gamma_0 = \beta_0$ and $k := 0$. At the kth iteration, for known x^k and γ_k. Find

$$\beta_{k+1} = \max\{\gamma_k, \mu\gamma_k, \mu^2\gamma_k, \ldots\}, x^{k+1} = P_C[x^k - \beta_{k+1}F(x^k)] \tag{3}$$

* Corresponding author.
E-mail addresses: ydong@zzu.edu.cn (Y. Dong), zhangxue2100442@163.com (X. Zhang).

0096-3003/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amc.2013.06.015

such that the following inequality

$$\beta_{k+1}\|F(x^k)-F(x^{k+1})\|^2 \leqslant (2-\delta)\langle x^k-x^{k+1}, F(x^k)-F(x^{k+1})\rangle \quad (4)$$

holds. Then set $\gamma_{k+1}=\beta_{k+1}$ or

$$\gamma_{k+1}=(1+\tau_{k+1})\beta_{k+1}. \quad (5)$$

To guarantee global convergence, He et al. [9] assumed that F is both Lipschitz continuous and strongly monotone. However, for such a self-adaptive projection method, its implementation needs neither Lipschitz constant nor strong monotonicity modulus.

Subsequently, Han and Sun [10] observed that the step length just described may hardly increase after certain number of iterations because the sequence $\{\tau_k\}$ is required to be summable for the method's convergence. So, they suggested replacing (4) and (5) by

$$(2-\delta)\beta_{k+1}\langle x^k-x^{k+1}, F(x^k)-F(x^{k+1})\rangle-\beta_{k+1}^2\|F(x^k)-F(x^{k+1})\|^2 \geqslant \max\left\{(\beta_{k+1}^2/\beta_k^2-1)\|x^k-P_C[x^k-\beta_kF(x^k)]\|^2, 0\right\} \quad (6)$$

and updating $\gamma_{k+1}=\beta_{k+1}$ or

$$\gamma_{k+1}=\min\{\beta_{\max}, \mu^{-1}\beta_{k+1}\}, \quad (7)$$

where $\beta_{\max}$ is a prescribed positive number larger than β_0. As shown in [10], the resulting self-adaptive projection method shares all nice properties of He's version [9]. Furthermore, the involved step length can enlarge during the whole iteration process (if necessary).

Motivated by their work, in this paper, we give a new modification of He's step length rule. We instead replace (4) and (5) by

$$\|x^k-P_C[x^k-\beta_kF(x^k)]\|^2 \geqslant \delta\beta_{k+1}\langle x^k-x^{k+1}, F(x^k)-F(x^{k+1})\rangle+\|x^k-x^{k+1}-\beta_{k+1}(F(x^k)-F(x^{k+1}))\|^2 \quad (8)$$

and select γ_{k+1} in the following way

$$\gamma_{k+1}:=\mu^p\beta_{k+1} \quad \text{with } p\in\{-1,0,1,\ldots\} \quad (9)$$

(See the third paragraph of Section 5 for a self-adaptive way of determining p.) Note that there are several common points between our modification and their work [9,10]. Firstly, at each iteration, every trial only needs one projection computation and one function evaluation. Secondly, the step length can be located by taking advantage of the information on two consecutive iterates, without evaluating Lipschitz constant and strong monotonicity modulus in advance.

Compared to the step length rule of He et al. [9] and its modification of Han and Sun [10], however, we would like to specially emphasize that our modification has an obvious advantage over them. It is just with possibly minimal parameters: δ, μ. In contrast, besides these two parameters, He et al. [9] introduced the τ-sequence whereas Han and Sun [10] posed an uniformly upper bound of the step lengths (see (7)). Such parameters are only for convergence in theory and are not beneficial to practical implementation of the individual self-adaptive projection method. More importantly, as explained in Remark 3.2 below, our modification still has a potential advantage of locating a acceptable step length after less times of trials in Armijo-type conditions.

The rest of this paper is organized as follows. In Section 2, we review some basic concepts and definitions used in this paper, and some important properties of the usual projection. In Section 3, we give a detailed description of the self-adaptive projection method suggested in this paper. In Section 4, we prove its global convergence at a R-linear rate provided that the underlying mapping is strongly monotone and Lipschitz continuous. Here we would like to point out that such a result on rate of convergence is absent in the literature [9,10]. In Section 5, preliminary numerical results show that it is flexible and effective in solving some test problems.

2. Preliminary results

In this section, we give a couple of useful concepts and some fundamental properties of the usual projection.

Definition 2.1. A mapping $F: R^n \to R^n$ is said to be L-Lipschitz continuous if there exists some positive number $L>0$ such that $\|F(x)-F(y)\| \leqslant L\|x-y\|$ for all $x, y \in R^n$.

Definition 2.2. A mapping $F: R^n \to R^n$ is said to be α-strongly monotone if there exists some positive number $\alpha>0$ such that $\langle x-y, F(x)-F(y)\rangle \geqslant \alpha\|x-y\|^2$ for all $x, y\in R^n$.

Let $P_C(x) := \text{argmin}\{\|x-y\| : y\in C\}$ be the usual projection onto some nonempty closed convex subset C of R^n. Then

$$\langle x-P_C(x), y-P_C(x)\rangle \leqslant 0, \quad \forall x\in R^n, \ \forall y\in C.$$

This is the most fundamental property of the usual projection. From this, one can arrive at the following two lemmas.